



Call for Papers

Mining the pharmacogenomics literature

We are inviting submissions for a special issue of the *Journal of Biomedical Informatics* on the automatic or semi-automatic extraction of relationships between biomedical entities relevant to pharmacogenomics from the research literature. Accepted papers will focus particularly on methods for the extraction of genotype–phenotype, genotype–drug, and phenotype–drug relationships and the novel use of these relationships for advancing pharmacogenomic research. Efforts aimed at creating benchmark corpora as well as comparative evaluation of existing relationship extraction methods are of special interest.

Pharmacogenomics is the study of how human genetic variation affects an individual's response to drugs. It includes a spectrum of discovery and application, ranging from the discovery of individual gene–drug interactions, the uncovering of pathways of drug response, the use of human and model system genetic variation to understand the mechanism of drug action, and the understanding of both the anticipated and unanticipated effects of drugs. The field includes the understanding of the role of gene–drug interactions in clinical medicine, including the identification of study cohorts, the definition of target populations, the evaluation of benefit, and the analysis of research and clinical databases to extract information about the larger context of drug interactions in different disease contexts. As such, pharmacogenomics is a timely and important field. The promise that it holds for individualized medicine is central as technical advances such as SNP microarrays, whole genome sequencing and other high-throughput measurement technologies allow us to predict beneficial, non-beneficial, and deleterious drugs for specific individuals based on aspects of both the individual and the drug.

However, information management in this field relies on fairly traditional means, especially curated databases (PharmGKB, DrugBank, dbGAP, and others), which do not scale to (1) the rapid expansion of the pharmacogenomics literature in recent years and (2) the increasingly available volume of full text publications, which contain more specific and (potentially) informative facts than Medline abstracts. Hence, although there is a large demand and significant utility of text analytics for the study of pharmacogenomics, the potential of such methods is not fully realized; in part because the work to date has failed to bridge the two distinct worlds—that of (bench) molecular biology and that of (clinically oriented) pharmacology—and because the developers of text analytics, both in computer science and biomedical informatics, are not fully aware of this challenging subfield.

The steady stream of work on extracting interactions from text, the increasing attention in the Semantic Web to capturing facts as "nano-publications" (individual assertions that are attributable to authors and traceable in their publications), and efforts to represent scientific discourse in a structured manner, all indicate that the time is ripe for research that goes even beyond the mere extraction of explicitly stated knowledge in documents to linking text-mined and database elements through formal reasoning to

uncover implicit and in some sense "new" knowledge. There have recently been scientific workshops and sessions at conferences devoted to text mining in the context of pharmacogenomics, including the 2010 and 2011 workshops of the Pacific Symposium on Biocomputing (<http://psb.stanford.edu/>), which have demonstrated the emerging critical mass of investigators in this subspecialty.

In order to advance this agenda, it is also essential that existing relationship extraction methods be compared to one another and that a community-wide sharable benchmark corpus emerges against which such efforts can be compared. We welcome submissions to this special issue that utilize information available at PharmGKB to compare different relationship extraction methods and the corresponding "new" knowledge discovery they might drive. These include curated relationships and annotations of genetic variants available at http://www.pharmgkb.org/resources/downloads_and_web_services.jsp.

The planned special issue aims to address the gap in coverage of text mining for pharmacogenomics, as an important initial application area of genomics in clinical medicine, and thus an important translational medicine activity. The technical area of the issue is intended to focus particularly on genotype–phenotype–drug relationships. It will include broad categories of work that have been well-studied in the past, specifically text mining and reasoning, but will restrict submissions to applications of that work to the constrained area of pharmacogenomics, and particularly genotype–phenotype–drug relationships. For example, topics that are solicited include:

- Relation extraction between genotypes, phenotypes, and drugs, and other semantic classes relevant to pharmacogenomics.
- Corpus development for pharmacogenomics text mining.
- Associating gene variants (mutations, alleles, rs/ss numbers) to the associated gene name.
- Using text mining to extract information about the association of drugs with clinical phenotypes.
- The use of biological networks in combination with text mining to facilitate discovery.
- Work on the corpus of documents linked to by PharmGKB.
- Reasoning systems applied over the PharmGKB knowledge base.
- The creation of ontologies to help relate molecular action of drugs to their clinical effects.

Work on gene taggers alone will not be considered responsive to this call. The key feature we seek in submissions is the use of language technologies to understand the molecular basis of drug response, its variability, and its impact on phenotypes at the molecular, cellular, organ and whole organisms level. Approaches that combine text-mining and knowledge-based systems are of special interest.

Peer review process

All submitted papers will go through a rigorous peer-review process that will include both programmatic relevance as well as scientific quality. All submissions should follow the guidelines for authors available at the *Journal of Biomedical Informatics* web site (<http://www.elsevier.com/locate/yjbin>). JBI's editorial policy is also outlined on that page and will be strictly followed by special issue reviewers. Note that JBI generally publishes at least one methodological review paper in each issue of the journal, and we would welcome a review of the state of the art in text mining for pharmacogenomics for this special issue.

Authors may contact the special issue lead editors (kevin.cohen@gmail.com or russ.altman@stanford.edu) with a proposed abstract before submission to get a sense for fit between the proposed manuscript and the goals of the special issue.

Submission process

Authors must submit their paper via the online Elsevier Editorial System (EES) at <http://ees.elsevier.com/jbi>. Authors can register and upload their text, tables, and figures, as well as subsequent re-

visions, through this website. Potential authors may contact the Publishing Services Coordinator in the journal's editorial office (jbi@elsevier.com) for questions regarding this process. When asked for the category of their submission, they should indicate that it is for the special issue on *Mining the pharmacogenomics literature* (which will appear as "Pharmacogenomics" on the pull-down menu).

Submission Deadline: March 15, 2011.

Guest Editors

K. Bretonnel Cohen
Russ Altman

Co-editors

Yael Garten
Udo Hahn
Nigam Shah